

# A sparsity augmented probabilistic collaborative representation based classification method

Journal of Algorithms &  
Computational Technology  
Volume 14: 1–10  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1748302620931042  
journals.sagepub.com/home/act



Xiao-Yun Cai<sup>1</sup> and He-Feng Yin<sup>2</sup>

## Abstract

In order to enhance the performance of image recognition, a sparsity augmented probabilistic collaborative representation based classification (SA-ProCRC) method is presented. The proposed method obtains the dense coefficient through ProCRC, then augments the dense coefficient with a sparse one, and the sparse coefficient is attained by the orthogonal matching pursuit (OMP) algorithm. In contrast to conventional methods which require explicit computation of the reconstruction residuals for each class, the proposed method employs the augmented coefficient and the label matrix of the training samples to classify the test sample. Experimental results indicate that the proposed method can achieve promising results for face and scene images. The source code of our proposed SA-ProCRC is accessible at <https://github.com/yinhefeng/SAProCRC>

## Keywords

Image recognition, probabilistic collaborative representation based classification, sparse representation, sparsity augmented

Received 9 October 2019; accepted 27 January 2019

## Introduction

Image recognition remains one of the hottest topics in the communities of computer vision and pattern recognition. During the past decade, sparse representation has been successfully applied in various domains. In face recognition, the pioneering work is the sparse representation based classification (SRC).<sup>1</sup> Concretely, SRC employs all the training samples as a dictionary, and a test sample is sparsely coded over the dictionary, then the classification is performed by checking which class yields the least reconstruction error. SRC can achieve promising recognition results even when the test samples are occluded or corrupted. To further promote the robustness of SRC, Wang et al.<sup>2</sup> proposed a correntropy matching pursuit (CMP) method for robust sparse representation based recognition. CMP can adaptively assign small weights on severely corrupted entries of data and large weights on clean ones, thus reducing the effect of large noise. Wu and Ding<sup>3</sup> presented a gradient direction-based hierarchical adaptive sparse and low-rank algorithm to tackle the real-world occluded face recognition problem. Gao et al.<sup>4</sup> developed a robust and discriminative low-rank representation method by exploiting the

low-rankness of both the data representation and each occlusion-induced error image simultaneously. Keinert et al.<sup>5</sup> designed a group sparse representation-based method for face recognition which introduces a non-convex sparsity-inducing penalty and a robust non-convex loss function.

Apart from classifier design, feature extraction is also a crucial stage in image recognition. The most classic subspace learning-based approaches are principal component analysis and linear discriminant analysis. Motivated by the recent development of sparse representation, Qiao et al.<sup>6</sup> presented a dimensionality reduction technique called sparsity preserving projections. To make SRC efficiently deal with high-dimensional data, Cui et al.<sup>7</sup> proposed an integrated

<sup>1</sup>School of Pharmacy and Medical Technology, Zhenjiang College, Zhenjiang, China

<sup>2</sup>School of Internet of Things Engineering, Jiangnan University, Wuxi, China

### Corresponding author:

Xiao-Yun Cai, Zhenjiang College, No. 518 Changxiangxi Avenue, College Park, Zhenjiang 212028, China.

Email: [dysfcxy@126.com](mailto:dysfcxy@126.com)



optimization algorithm to implement feature extraction, dictionary learning, and classification simultaneously. To tackle the corrupted data, Xie et al.<sup>8</sup> explored a dimensionality reduction method termed low-rank sparse preserving projections by combining the manifold learning and low-rank sparse representation.

Recently, sparse representation has been applied to a wide range of tasks. Zhang et al.<sup>9</sup> developed a structural sparse representation model for visual tracking. Liu et al.<sup>10</sup> introduced the convolutional sparse representation into image fusion. Guo et al.<sup>11</sup> proposed a sparse and dense hybrid representation-based target detector for hyperspectral imagery (HSI).

Another critical issue in sparse representation is how to solve the  $\ell_1$ -norm constraint problem. Zhang et al.<sup>12</sup> presented a survey of sparse representation algorithms and found that Homotopy and augmented Lagrange multiplier (ALM) can achieve better recognition performance and have relatively lower computational cost.

Akhtar et al.<sup>13</sup> revealed that sparseness explicitly contributes to improved classification. And they proposed a sparsity augmented collaborative representation based classification (SA-CRC) which employs both dense and sparse collaborative representations to recognize a test sample. However, collaborative representation based classification (CRC)<sup>14</sup> utilizes all the training samples to represent the input test sample, which neglects the relationship between the test sample and each of the multiple classes. To overcome the drawback of SA-CRC, first we obtain a dense representation by probabilistic collaborative representation based classification (ProCRC),<sup>15</sup> then we augment the representation of ProCRC with a sparse representation to further promote the sparsity of ProCRC. Moreover, different from conventional representation based classification methods that use class-wise reconstruction error for classification, we utilize the label matrix of training data and the augmented coefficient of a test sample for final classification. The proposed method is termed as sparsity augmented probabilistic collaborative representation based classification (SA-ProCRC). In summary, our contributions are as follows:

- We promote the sparsity of ProCRC by augmenting the representation of ProCRC with a sparse representation.
- We employ an efficient classification rule to recognize the test sample, in which the explicit computation of residuals class by class is avoided.
- Experimental results on diverse datasets validate the efficacy of our proposed method.

## Related work

Given  $n$  training samples belonging to  $C$  classes, and the training data matrix is denoted by  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_C] = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ , where  $\mathbf{X}_i$  is the data matrix of the  $i$ -th class. The  $i$ -th class has  $n_i$  training samples and  $\sum_{i=1}^C n_i = n$ ,  $i = 1, 2, \dots, C$ ,  $m$  is the dimensionality of vectorized samples.

### Sparse representation based classification

In SRC,<sup>1</sup> a test sample  $\mathbf{y} \in \mathbb{R}^m$  is first represented as a sparse linear combination of all the training data, then the classification is performed by checking which class leads to the least reconstruction error, the objective function of SRC is formulated as

$$\min_{\alpha} \|\alpha\|_1, \text{ s.t. } \|\mathbf{y} - \mathbf{X}\alpha\|_2^2 \leq \varepsilon \quad (1)$$

where  $\varepsilon$  is a given error tolerance. When we obtain the coefficient vector  $\alpha$  of  $\mathbf{y}$ , the test sample  $\mathbf{y}$  is classified according to the following formulation

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{y} - \mathbf{X}_i \alpha_i\|_2 \quad (2)$$

where  $\alpha_i$  is the coefficient vector that corresponds to the  $i$ -th class.

### Collaborative representation based classification

SRC and its extensions have achieved encouraging results in a variety of pattern classification tasks. However, Zhang et al.<sup>14</sup> argued that it is the collaborative representation mechanism rather than the  $\ell_1$ -norm sparsity that makes SRC powerful for classification. And they presented CRC algorithm, which replaces the  $\ell_1$ -norm in SRC with the  $\ell_2$ -norm constraint, the objective function of CRC is formulated as follows

$$\min_{\alpha} \|\mathbf{y} - \mathbf{X}\alpha\|_2^2 + \lambda \|\alpha\|_2^2 \quad (3)$$

CRC has the following closed-form solution

$$\alpha = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (4)$$

where  $\mathbf{I}$  is the identity matrix. Let  $\mathbf{P} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T$ , one can see that  $\mathbf{P}$  is determined by the training data matrix  $\mathbf{X}$ . Therefore, when given all the training data,  $\mathbf{P}$  can be pre-computed, which makes CRC very efficient. CRC employs the following regularized residual for classification

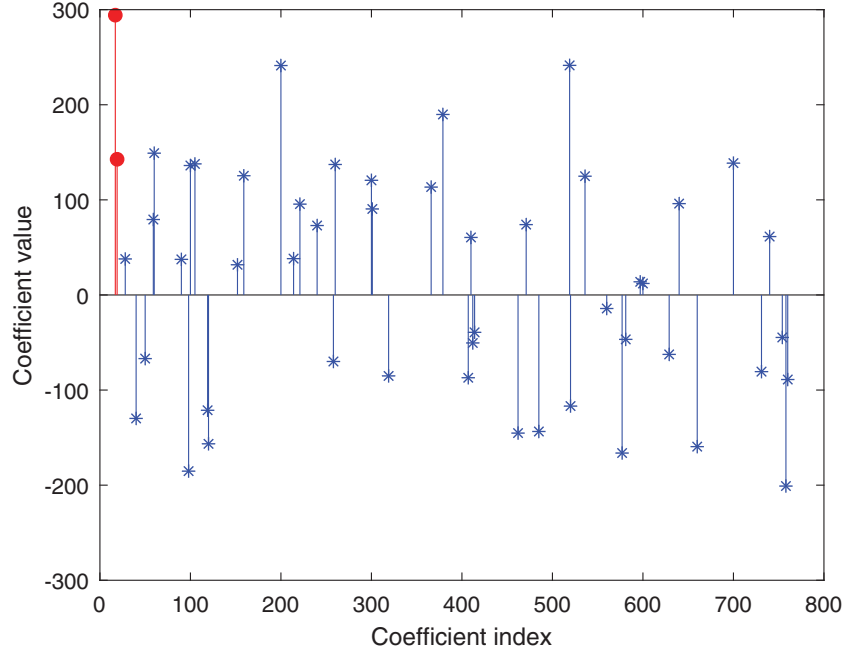
$$\text{identity}(y) = \arg \min_i \frac{\|y - \mathbf{X}_i \alpha_i\|_2}{\|\alpha_i\|_2} \quad (5)$$

mechanism of CRC from a probabilistic perspective and developed a probabilistic collaborative representation based classifier (ProCRC), and the objective function of ProCRC is formulated as

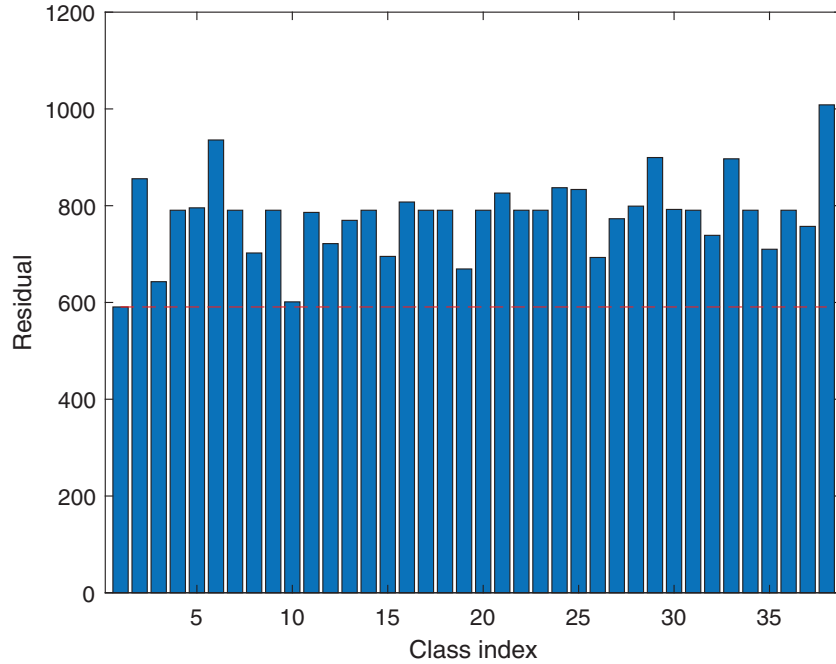
### Probabilistic CRC

Inspired by the work of probabilistic subspace approaches, Cai et al.<sup>15</sup> explored the classification

$$\min_{\tilde{\alpha}} \|y - \mathbf{X}\tilde{\alpha}\|_2^2 + \lambda \|\tilde{\alpha}\|_2^2 + \frac{\gamma}{C} \sum_{i=1}^C \|\mathbf{X}\tilde{\alpha} - \mathbf{X}_i \tilde{\alpha}_i\|_2^2 \quad (6)$$



**Figure 1.** Coefficients obtained by SRC.

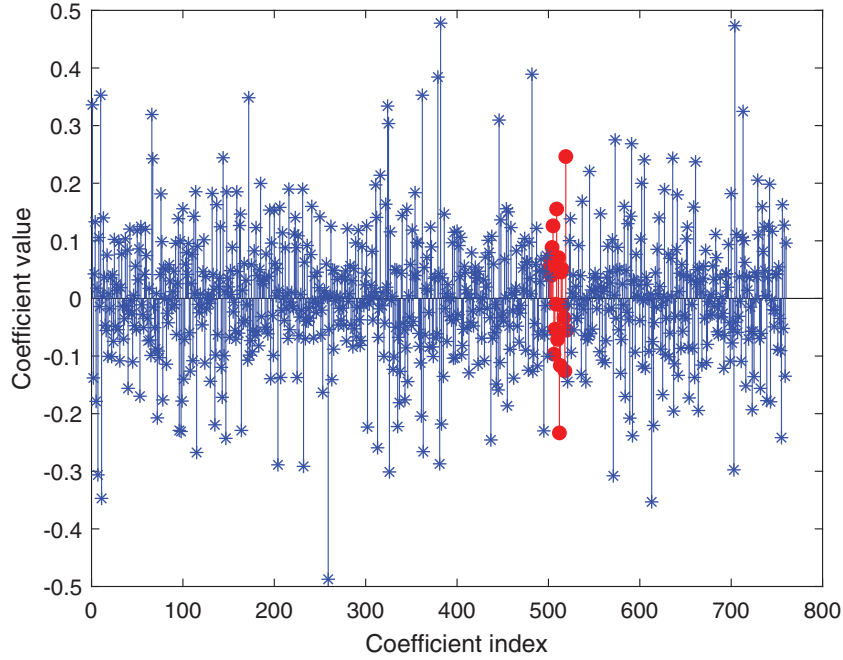


**Figure 2.** The residual of SRC for each class, and the first class has the least residual.

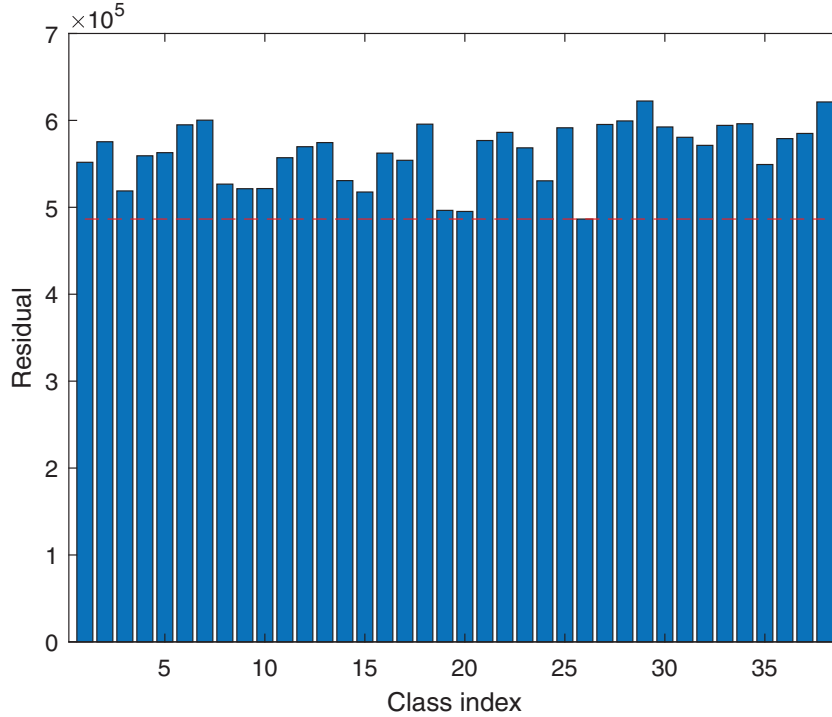
where  $\lambda$  and  $\gamma$  are two balancing parameters. One can see that ProCRC is reduced to CRC when  $\gamma=0$ . Suppose  $\mathbf{X}'_i$  is a matrix that has the same size as  $\mathbf{X}$ , and  $\mathbf{X}'_i$  only contains the samples from the  $i$ -th class, namely  $\mathbf{X}'_i = [\mathbf{0}, \dots, \mathbf{X}_i, \dots, \mathbf{0}]$ . Let  $\tilde{\mathbf{X}}'_i = \mathbf{X} - \mathbf{X}'_i$ , after

some deductions, we can obtain the following closed-form solution to ProCRC

$$\tilde{\alpha} = \mathbf{T}y \quad (7)$$



**Figure 3.** Coefficients computed by ProCRC.



**Figure 4.** The residual of ProCRC for each class, one can see that the 26th class has the minimal residual.

where  $\mathbf{T} = (\mathbf{X}^T \mathbf{X} + \frac{\lambda}{C} \sum_{i=1}^C (\bar{\mathbf{X}}'_i)^T \bar{\mathbf{X}}'_i + \lambda \mathbf{I})^{-1} \mathbf{X}^T$  and  $\mathbf{I}$  is the identity matrix.

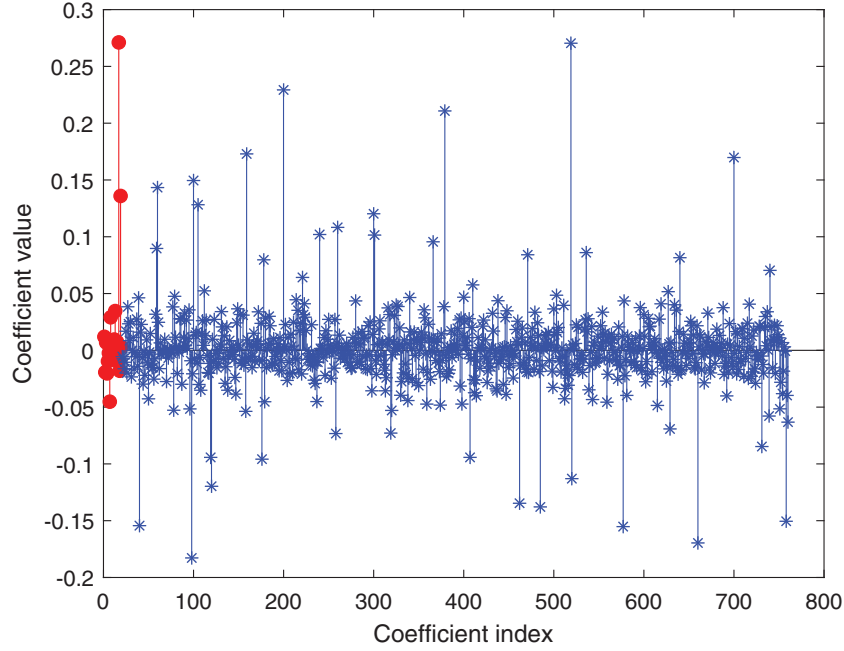
computed by OMP,<sup>16</sup> and the optimization problem for sparse representation is given by

$$\min_{\alpha} \|\mathbf{y} - \mathbf{X} \alpha\|_2, \text{ s.t. } \|\alpha\|_0 \leq k \quad (8)$$

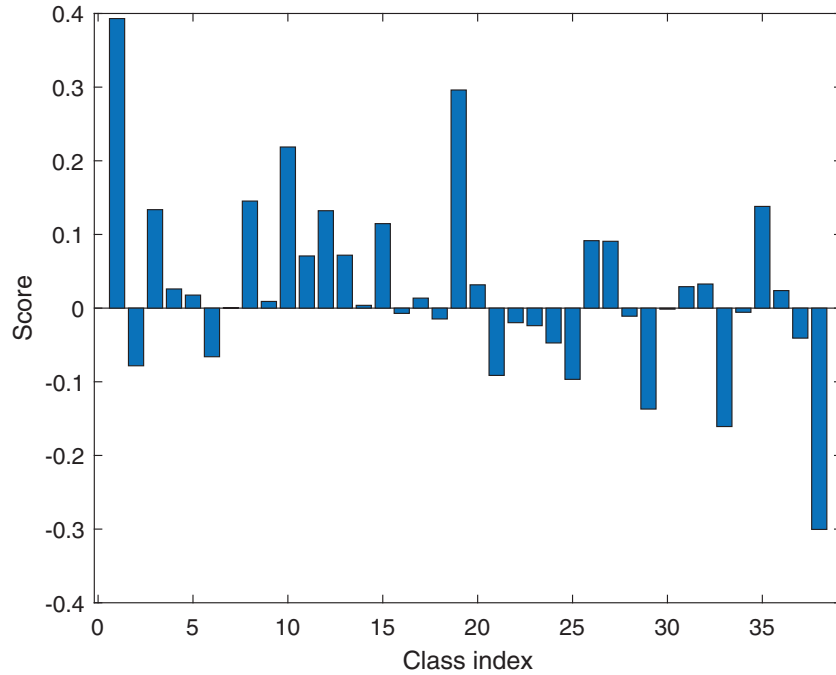
### SA-ProCRC

In our proposed SA-ProCRC, the dense representation of ProCRC is augmented by a sparse representation

where  $k$  is the sparsity level.



**Figure 5.** Coefficients obtained by SA-ProCRC.



**Figure 6.** The score of SA-ProCRC for each class, it is evident that the first class has the largest value.

The augmented coefficient  $\hat{\alpha}$  can be obtained according to the following formulation

$$\hat{\alpha} = \frac{\circ\alpha + \tilde{\alpha}}{\|\circ\alpha + \tilde{\alpha}\|_2} \quad (9)$$

where  $\circ\alpha$  is the sparse coefficient computed by OMP, and  $\tilde{\alpha}$  is the coefficient obtained by ProCRC.

Let  $\mathbf{L} = [\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_n] \in \mathbb{R}^{C \times n}$  be the label matrix of the training data, and  $\mathbf{l}_j = [0, 0, \dots, 1, \dots, 0]^T \in \mathbb{R}^{C \times 1}$  denotes the label vector of the  $j$ -th training sample. For the  $i$ -th class,  $\mathbf{L}$  consists of  $n_i$  non-zero elements in its  $i$ -th row, at the indices associated with the columns of  $\mathbf{X}_i$ . Remember that  $\mathbf{X}_i$  is the subset of dictionary atoms belonging to the  $i$ -th class. Therefore, the  $i$ -th entry of the vector  $\mathbf{q} = \mathbf{L} \hat{\alpha}$  expresses the sum of coefficients in  $\hat{\alpha}$  which correspond to the atoms in  $\mathbf{X}_i$ , and  $\mathbf{q}$  is dubbed as the score of each class. Consequently, the test sample is designated into the class that leads to the largest score.

Our proposed SA-ProCRC has the following procedures. First, the dense coefficient and sparse coefficient are obtained by solving equations (6) and (8), respectively. Second, the dense coefficient is augmented by the sparse coefficient. Finally, the test sample is recognized according to the augmented coefficient vector and the label matrix of the training data. Algorithm 1 presents our proposed scheme.

#### Algorithm 1. SA-ProCRC

**Table 1.** Details of datasets used in our experiments.

Dataset	# Sample	# Class	# Feature
Yale	165	15	576
EYaleB	2414	38	504
AR	2600	100	540
Scene 15	4485	15	3000

Note: The columns from left to right are the names of datasets, total number of samples, number of classes, and the dimensionality of features.

**Input:** Training data matrix  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_C] \in \mathbb{R}^{m \times n}$  and label matrix  $\mathbf{L}$ , test data  $\mathbf{y} \in \mathbb{R}^m$ , parameters  $\lambda$  and  $\gamma$  for ProCRC, sparsity level  $k$  for SRC.

**Output:**  $\text{label}(\mathbf{y}) = \arg\max_i(q_i)$

1. Compute the coefficient  $\tilde{\alpha}$  of ProCRC by using equation (7)
2. Obtain the sparse coefficient  $\circ\alpha$  of SRC by solving equation (8)
3. Compute the augmented coefficient  $\hat{\alpha} = \frac{\circ\alpha + \tilde{\alpha}}{\|\circ\alpha + \tilde{\alpha}\|_2}$
4. Compute  $\mathbf{q} = \mathbf{L} \hat{\alpha}$

#### Analysis of SA-ProCRC

In this section, we present some experimental results on the Extended Yale B database to illustrate the effectiveness of SA-ProCRC. The Extended Yale B database contains 38 individuals and there are about 64 images for each individual. We randomly select 20 images per subject as the training data; therefore, the dictionary contains 760 atoms. We select a test image which belongs to the first subject, and the sparse coefficients and corresponding residual for each class are plotted in Figures 1 and 2. It can be seen from Figure 1 that coefficients belong to the first class are prominent. From Figure 2, we can clearly see that the first class has the least residual, which indicates that the test sample is correctly classified by SRC. Figure 3 shows the coefficients derived by ProCRC, and we can see that the coefficients are rather dense. Figure 4 presents the residual of ProCRC, one can see that the 26th class has the least residual, thus the test sample is wrongly classified to the 26th class. Coefficients obtained by SA-ProCRC are shown in Figure 5, and we can see that coefficients from the first class are dominant. Figure 6 plots the score of SA-ProCRC for each class, it can be seen that the first class delivers the largest value. As a result, the test sample is designated to the first class by SA-ProCRC. From the above experimental results, we can find that the dense representation of ProCRC



**Figure 7.** Example images from the Yale database.



may lead to misclassification. By augmenting the dense representation with a sparse representation, the misclassification can be alleviated. This validates the superiority of our proposed SA-ProCRC.

## Experiments

In this section, we conduct experiments on four benchmark datasets: the Yale database, the Extended Yale B database, the AR database, and the Scene 15 dataset, the details of these datasets are listed in Table 1. We compare the proposed method with state-of-the-art representation based classification methods and several dictionary learning approaches, such as SRC,<sup>1</sup> CRC,<sup>14</sup> ProCRC,<sup>15</sup> discriminative K-SVD (D-KSVD),<sup>17</sup> label

consistent K-SVD (LC-KSVD),<sup>18</sup> fisher discrimination dictionary learning (FDDL),<sup>19</sup> dictionary learning based on commonalities and particularities (COPAR),<sup>20</sup> joint discriminative Bayesian dictionary and classifier learning (JBDC),<sup>21</sup> and SA-CRC.<sup>13</sup> For SRC, we solve the problem in equation (1) as in Wright et al.<sup>1</sup> For CRC, LC-KSVD, FDDL, COPAR, JBDC, and SA-CRC, we use the publicly available codes. We adapted the code of LC-KSVD to implement D-KSVD. For SA-CRC and our proposed SA-ProCRC, OMP is utilized to obtain the sparse representation. We utilize the same value of sparsity level ( $k = 50$ ) as in SA-CRC.<sup>13</sup> All experiments are run with MATLAB R2019a under Windows 10 on PC equipped with 3.60 GHz CPU and 16 GB RAM.

**Table 2.** Recognition accuracy on the Yale database.

Methods	Accuracy (%)
SRC	$95.06 \pm 3.32$
CRC	$94.53 \pm 2.97$
ProCRC	$95.33 \pm 2.82$
D-KSVD	$94.26 \pm 2.88$
LC-KSVD	$94.53 \pm 0.03$
FDDL	$95.73 \pm 3.00$
COPAR	$91.33 \pm 4.23$
JBDC	$94.93 \pm 2.72$
SA-CRC	$95.60 \pm 2.59$
SA-ProCRC	<b><math>96.13 \pm 2.84</math></b>

SRC: sparse representation based classification; CRC: collaborative representation based classification; ProCRC: probabilistic collaborative representation based classification; SA-CRC: sparsity augmented collaborative representation based classification; SA-ProCRC: sparsity augmented probabilistic collaborative representation based classification. Bold value signifies the best recognition accuracy.

### Experiments on the Yale database

There are 165 images for 15 subjects in the Yale database, each has 11 images. These images have illumination and expression variations, Figure 7 shows some example images from this database. All the images are resized to  $24 \times 24$  pixels, leading to a 576-dimensional vector. In our experiments, six images per subject are randomly selected for training and the rest for testing. The error tolerance  $\varepsilon$  of SRC is 0.05, and the balancing parameter  $\lambda$  of CRC is 0.001. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 30 and 60, respectively. Sparsity level  $k$  and  $\lambda$  of SA-CRC are set to be 50 and 0.002, respectively. Experimental results are summarized in Table 2, in which the best result is highlighted by bold number. It can be observed that SA-ProCRC achieves the highest recognition accuracy, with a 17% reduction in the error rate of ProCRC, and 12% reduction in that of SA-CRC.



**Figure 8.** Example images from the Extended Yale B database.

### Experiments on the Extended Yale B database

The Extended Yale B face database is composed of 2414 images of 38 individuals. Each individual has 59–64 images taken under different illumination conditions, example images from this dataset are shown in Figure 8. In our experiments, each  $192 \times 168$  image is projected onto a 504-dimensional space via random

projection. Twenty images per person are selected for training and the remaining for testing. We use the error tolerance of 0.05 for SRC, and the regularization parameter  $\lambda = 0.001$  for CRC. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 50 and 400, respectively. Sparsity level  $k$  and  $\lambda$  of SA-CRC are set to be 50 and 0.005, respectively. Table 3 lists the recognition accuracy of the comparison methods. It

**Table 3.** Recognition accuracy on the Extended Yale B database.

Methods	Accuracy (%)
SRC	$93.18 \pm 0.55$
CRC	$94.77 \pm 0.48$
ProCRC	$94.82 \pm 0.49$
D-KSVD	$90.79 \pm 0.51$
LC-KSVD	$91.48 \pm 0.69$
FDDL	$92.32 \pm 0.68$
COPAR	$90.81 \pm 0.55$
JBDC	$94.74 \pm 0.83$
SA-CRC	$95.52 \pm 0.73$
SA-ProCRC	<b><math>95.64 \pm 0.78</math></b>

SRC: sparse representation based classification; CRC: collaborative representation based classification; ProCRC: probabilistic collaborative representation based classification; SA-CRC: sparsity augmented collaborative representation based classification; SA-ProCRC: sparsity augmented probabilistic collaborative representation based classification. Bold value signifies the best recognition accuracy.

**Table 4.** Recognition accuracy on the AR database.

Methods	Accuracy (%)
SRC	$91.25 \pm 1.17$
CRC	$92.04 \pm 0.83$
ProCRC	$93.03 \pm 0.64$
D-KSVD	$90.31 \pm 1.13$
LC-KSVD	$89.31 \pm 1.27$
FDDL	$91.01 \pm 0.99$
COPAR	$89.06 \pm 1.54$
JBDC	$90.97 \pm 0.79$
SA-CRC	$93.74 \pm 0.84$
SA-ProCRC	<b><math>94.67 \pm 0.66</math></b>

SRC: sparse representation based classification; CRC: collaborative representation based classification; ProCRC: probabilistic collaborative representation based classification; SA-CRC: sparsity augmented collaborative representation based classification; SA-ProCRC: sparsity augmented probabilistic collaborative representation based classification. Bold value signifies the best recognition accuracy.



**Figure 9.** Example images from the AR database.





**Figure 10.** Example images from the Scene 15 dataset.

can be seen that our proposed SA-ProCRC is superior to its competing approaches.

### Experiments on the AR database

The AR database has more than 4000 face images of 126 subjects with variations in facial expression, illumination conditions, and occlusions, Figure 9 shows example images from this database. We use a subset of 2600 images of 50 male and 50 female subjects from the database. Each  $165 \times 120$  face image is projected onto a 540-dimensional vector by random projection. Ten images per person are randomly selected for training and the remaining for testing. The error tolerance of SRC is 0.05, and the balancing parameter of CRC is 0.0014. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 50 and 600, respectively. Sparsity level  $k$  and  $\lambda$  of SA-CRC are set to be 50 and 0.002, respectively. Experimental results are shown in Table 4. We can see that the best classification result is achieved by our proposed SA-ProCRC, with a 23% reduction in the error rate of ProCRC.

### Experiments on the Scene 15 dataset

This dataset contains 15 natural scene categories including a wide range of indoor and outdoor scenes, such as bedroom, office, and mountain, example images from this dataset are shown in Figure 10. For fair comparison, we employ the 3000-dimensional scale invariant feature transform (SIFT)-based features used in LC-KSVD.<sup>18</sup> We randomly select 50 images per category as training data and use the rest for testing. The error tolerance of SRC is  $1e-6$ , and the balancing parameter of CRC is 1. Fifty atoms are used for D-KSVD and LC-KSVD. Sparsity level  $k$  and  $\lambda$  of SA-CRC are set to be 50 and 1, respectively. Recognition

**Table 5.** Recognition accuracy on the Scene 15 dataset.

Methods	Accuracy (%)
SRC	$95.41 \pm 0.13$
CRC	$96.15 \pm 0.33$
ProCRC	$96.56 \pm 0.35$
D-KSVD	$95.12 \pm 0.18$
LC-KSVD	$96.37 \pm 0.28$
FDDL	$94.08 \pm 0.43$
COPAR	$96.02 \pm 0.28$
JBDC	$97.36 \pm 0.32$
SA-CRC	$97.18 \pm 0.25$
<b>SA-ProCRC</b>	<b><math>97.56 \pm 0.20</math></b>

SRC: sparse representation based classification; CRC: collaborative representation based classification; ProCRC: probabilistic collaborative representation based classification; SA-CRC: sparsity augmented collaborative representation based classification; SA-ProCRC: sparsity augmented probabilistic collaborative representation based classification. Bold value signifies the best recognition accuracy.

accuracy of different approaches on this dataset is presented in Table 5. Again, SA-ProCRC outperforms the comparison methods.

## Conclusions

It has been argued that it is the collaborative representation mechanism rather than the sparsity constraint that makes SRC powerful for pattern classification. As a result, sparsity is ignored to some extent in CRC and its extensions. To address this problem, we present a SA-ProCRC method to promote the sparsity in ProCRC. The proposed SA-ProCRC is

computationally efficient due to the fact that ProCRC has closed-form solution. Meanwhile, discriminative information containing in the resulting sparse coefficient can be exploited in SA-ProCRC. In essence, SA-ProCRC is a classifier, thus it can be applied to other pattern classification tasks. In our future work, we will evaluate SA-ProCRC with deep features and develop new representation based classification algorithm.

### Acknowledgements

The authors would like to thank Prof. Naveed Akhtar for providing the source code of SA-CRC at <http://staffhome.ecm.uwa.edu.au/~00053650/code.html>

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (Grant No. 61672265).

### ORCID iDs

Xiao-Yun Cai  <https://orcid.org/0000-0002-0880-3938>

He-Feng Yin  <https://orcid.org/0000-0001-5831-1475>

### References

1. Wright J, Yang AY, Ganesh A, et al. Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 2009; 31: 210–227.
2. Wang Y, Tang YY and Li L. Correntropy matching pursuit with application to robust digit and face recognition. *IEEE Trans Cybern* 2016; 47: 1354–1366.
3. Wu CY and Ding JJ. Occluded face recognition using low-rank regression with generalized gradient direction. *Pattern Recognit* 2018; 80: 256–268.
4. Gao G, Yang J, Jing XY, et al. Learning robust and discriminative low-rank representations for face recognition with occlusion. *Pattern Recognit* 2017; 66: 129–143.
5. Keinert F, Lazzaro D and Morigi S. A Robust group-sparse representation variational method with applications to face recognition. *IEEE Trans Image Process* 2019; 28: 2785–2798.
6. Qiao L, Chen S and Tan X. Sparsity preserving projections with applications to face recognition. *Pattern Recognit* 2010; 43: 331–341.
7. Cui Y, Jiang J, Lai Z, et al. An integrated optimisation algorithm for feature extraction, dictionary learning and classification. *Neurocomputing* 2018; 275: 2740–2751.
8. Xie L, Yin M, Yin X, et al. Low-rank sparse preserving projections for dimensionality reduction. *IEEE Trans Image Process* 2018; 27: 5261–5274.
9. Zhang T, Xu C and Yang MH. Robust structural sparse tracking. *IEEE Trans Pattern Anal Mach Intell* 2018; 41: 473–486.
10. Liu Y, Chen X, Ward RK, et al. Image fusion with convolutional sparse representation. *IEEE Signal Process Lett* 2016; 23: 1882–1886.
11. Guo T, Luo F, Zhang L, et al. Target detection in hyperspectral imagery via sparse and dense hybrid representation. *IEEE Geosci Remote Sensing Lett* 2019; 17: 716–720.
12. Zhang Z, Xu Y, Yang J, et al. A survey of sparse representation: algorithms and applications. *IEEE Access* 2015; 3: 490–530.
13. Akhtar N, Shafait F and Mian A. Efficient classification with sparsity augmented collaborative representation. *Pattern Recognit* 2017; 65: 136–145.
14. Zhang L, Yang M and Feng X. Sparse representation or collaborative representation: which helps face recognition? In: *CVPR*, Colorado Springs, CO, 20–25 June 2011, pp.471–478.
15. Cai S, Zhang L, Zuo W, et al. A probabilistic collaborative representation based approach for pattern classification. In: *CVPR*, Las Vegas, NV, 27–30 June 2016, pp.2950–2959.
16. Tropp JA and Gilbert AC. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans Inform Theory* 2007; 53: 4655–4666.
17. Zhang Q and Li B. Discriminative K-SVD for dictionary learning in face recognition. In: *CVPR*, San Francisco, CA, 13–18 June 2010, pp.2691–2698.
18. Jiang Z, Lin Z and Davis LS. Learning a discriminative dictionary for sparse coding via label consistent K-SVD. In: *CVPR*, Colorado Springs, CO, 20–25 June 2011, pp.697–1704.
19. Yang M, Zhang L, Feng X, et al. Fisher discrimination dictionary learning for sparse representation. In: *CVPR*, Colorado Springs, CO, 20–25 June 2011, pp.543–550.
20. Kong S and Wang D. A dictionary learning approach for classification: separating the particularity and the commonality. In: *ECCV*, Florence, Italy, 7–13 October 2012, pp.186–199.
21. Akhtar N, Mian A and Porikli F. Joint discriminative Bayesian dictionary and classifier learning. In: *CVPR*, Honolulu, HI, 21–26 July 2017, pp.1193–1202.